Understanding the multimodal nature of deictic reference

Gozdem Arikan[1], Peter Boddy, Kenny R. Coventry[1]

[1] University of East Anglia

g.arikan@uea.ac.uk, k.coventry@uea.ac.uk

Spatial demonstratives (in English *this* and *that*), gesture and eye gaze (deictic communication) are used in synchrony to establish joint attention and draw attention of interlocutors to a specific position in space (Bangerter, 2004; Diessel, 2006, 2014; Stukenbrock, 2020; Todisco et al., 2020). However, the relative importance of these cues for deictic reference and the role of eye gaze and gesture in determining demonstrative choice have not thus far been illuminated (Bangerter, 2004; Cooperrider, 2016; Coventry et al., 2014). In three online experiments (Gorilla inc.), we manipulated the congruency of deictic cues to establish their relative importance. In Experiment 1, we observe whether the use of gesture and eye gaze are necessary to produce demonstratives or whether knowledge of object location is enough to choose a demonstrative form. In the Experiments 2 and 3, we tested which one of the deictic cues, demonstratives, pointing and eye gaze, is *dominant* in guiding attention, and whether there is an effect of stimuli type (image or video clips) on response. Participants saw images or videos of an individual (agent; male, him) referring to one of the two items on a table, placed proximally or distally from himself by pointing (left hand only) and gazing towards the target (Figure 1). The images were coupled either with a noun (as a cue for object location) labelling one of the two items (Experiment 1) or with a spatial demonstrative either in written format (Experiment 2) or as audio (Experiment 3). The study population consisted of Psychology students and participants from Prolific recruitment platform, all native speakers of English.

Experiment 1 (*N*=38): In this experiment, participants were asked to pick a spatial demonstrative they thought the agent would utter based on the object location (noun), pointing and eye gaze information in order to complete a two-word sentence (Figure 1a). The experiment had a 2 (object location: proximal, distal) x3 (pointing: proximal, distal, no pointing) x3 (eye gaze: proximal, distal, eyes closed) repeated measures design. The results suggested the strongest cue for picking a spatial demonstrative (word *this* in particular) was the object location of the referent, followed by pointing. There was also a significant interaction between object location and pointing. While seeing a pointing arm as an addition to knowledge on object location did not affect demonstrative choice compared to knowing the object location only, it was the case that a mismatch between object location and pointing led to an increased choice of *that*.

Experiment 2 (*N*=33): Experiment 2 examined the effect of demonstrative, eye gaze and pointing on the choice of referent participants thought the speaker was referring to (near object or far object) (Figure 1b.). Based on the deictic cues seen, the participants chose between one of the two potential referents they thought the agent would choose. The results suggested that spatial demonstratives, pointing and eye gaze are all important deictic cues to object reference. However, there were a number of significant interactions between pointing and other variables. For example, seeing a pointing arm and demonstrative (congruently) increased the decisiveness of referent choice both proximally and distally. Furthermore, when pointing and demonstrative did not match, pointing was the dominant cue to referent choice. Moreover, eye gaze only affected referent choice in the absence of pointing.

Experiment 3 (*N*=73): The purpose of this experiment was to replicate Experiment 2 with addition of a new level to the design; stimuli type. The participants saw videos (dynamic) and images (static) as stimuli. Additionally, the demonstratives were now in audio format for both images and videos (Figure 1c.). The findings of this experiment mainly replicated Experiment 2, but with some additional findings. All deictic cues and the stimuli type were significant factors. The results from trials with static images replicated the Experiment 2 with an additional interaction between demonstrative and eye gaze. The interaction of demonstratives and eye gaze did not suggest an advantage of combining eye gaze with spatial demonstratives, but there was a cost to incongruence between demonstratives and eye gaze. Trials with videos as stimuli also had the same trend as in Experiment 2 but there was an additional a three-way interaction of the deictic cues. The three-way interaction overall suggested that pointing is the most dominant cue and in absence of pointing, spatial demonstratives and eye gaze have an effect.

Conclusion: When choosing a demonstrative, knowledge of the object location is the most important information. While eye gaze might not have any function in this dichotomous choice, pointing is the next factor which influence the demonstrative one might utter (Experiment 1). Amongst the three deictic cues (spatial demonstrative, pointing and eye gaze), pointing is the strongest cue, followed by eye gaze and spatial demonstrative seen (Experiment 2 and 3).

Index Terms: spatial demonstratives, pointing, eye gaze.

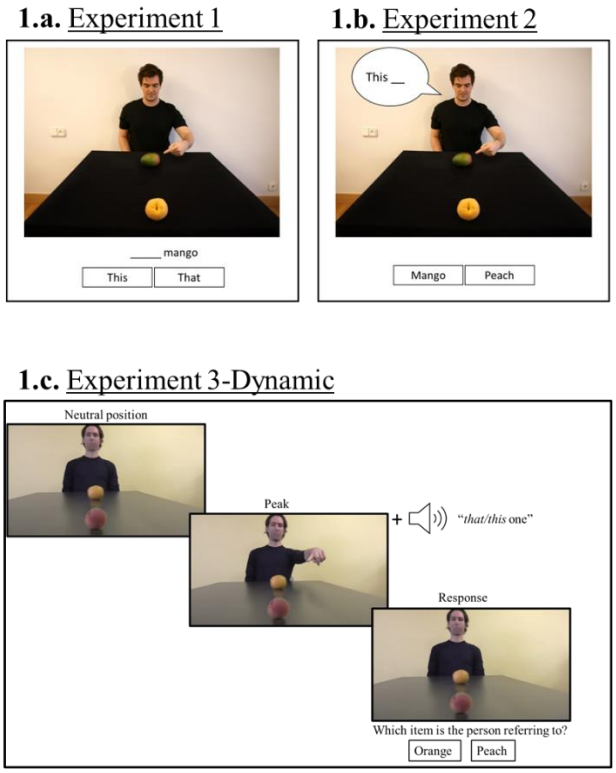

Figure 1: *Stimuli example from Experiment 1, 2 and 3.*

References

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychological Science*, *15*(6), 415–419. https://doi.org/10.1111/j.0956-7976.2004.00694.x

Cooperrider, K. (2016). The Co-Organization of Demonstratives and Pointing Gestures. *Discourse Processes*, *53*(8), 632–656. https://doi.org/10.1080/0163853X.2015.1094280

Coventry, K. R., Griffiths, D., & Hamilton, C. J. (2014). Spatial demonstratives and perceptual space: Describing and remembering object location. *Cognitive Psychology*, *69*, 46–70. https://doi.org/10.1016/j.cogpsych.2013.12.001

Diessel, H. (2006). Demonstratives, joint attention, and the emergence of grammar. *Cognitive Linguistics*, *17*(4), 463–489. https://doi.org/10.1515/COG.2006.015/MACHINEREADABLECITATION/RIS

Diessel, H. (2014). Demonstratives, frames of reference, and semantic universals of space. *Linguistics and Language Compass*, *8*(3), 116–132. https://doi.org/10.1111/lnc3.12066

Stukenbrock, A. (2020). Deixis, Meta-Perceptive Gaze Practices, and the Interactional Achievement of Joint Attention. *Frontiers in Psychology*, *11*, 1779. https://doi.org/10.3389/FPSYG.2020.01779

Todisco, E., Guijarro-Fuentes, P., Collier, J., & Coventry, K. R. (2020). The temporal dynamics of deictic communication. *First Language*, 014272372093678. https://doi.org/10.1177/0142723720936789